

DOCUMENT RESUME

ED 465 544

SE 066 199

AUTHOR Haury, David L.
TITLE The State of State Proficiency Testing in Science. ERIC Digest.
INSTITUTION ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, OH.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO EDO-SE-01-10
PUB DATE 2001-12-00
NOTE 4p.
CONTRACT ED-99-CO-0024
AVAILABLE FROM ERIC Clearinghouse for Science, Mathematics, and Environmental Education, 1929 Kenny Road, Columbus, OH 43210-1080. Tel: 800-276-0462 (Toll Free); Fax: 614-292-0263. For full text: <http://www.ericse.org>.
PUB TYPE ERIC Publications (071) -- ERIC Digests in Full Text (073)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Tests; Elementary Secondary Education; *National Standards; Science Instruction; Science Teachers; *State Standards; *Testing
IDENTIFIERS ERIC Digests

ABSTRACT

Schools across the United States are striving to improve student performance in science by adjusting curricula and teaching practices to meet national and state standards. This ERIC Digest describes the current status of proficiency testing in science, including the topics of high stakes testing and its outcomes. It also discusses the future of science as it is affected by the high level of testing. As states quickly begin to align their science testing with the current standards, they must remember to have their teachers "teach to the standards." (MVL)

The State of State Proficiency Testing in Science

ERIC Digest

David L. Haury

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☐ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.
- ☐ Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Schools across the United States are striving to improve student performance in science by adjusting curricula and teaching practices to meet national and state standards. *Standards-based reform* is the rallying cry for these efforts to enliven the *National Science Education Standards* (NSES; National Research Council, 1996). Ongoing reform in science education has intensified in response to the results of widely reported national and international studies of student understanding. Despite rapid advancements in science and technology within the nation, most U.S. school students have not performed all that well on tests of scientific knowledge and understanding.

The most recent results in science from the National Assessment of Educational Progress show no statistically significant changes in average student scores at grades 4 or 8 since 1996, but the average scores for students in grade 12 have declined (See <http://nces.ed.gov/nationsreportcard/science/results/>). Results from the Third International Mathematics and Science Study (TIMSS) were even more jarring. Though results across the states were highly variable, U.S. students overall achieved mediocre scores compared to the students of other developed nations (U.S. National TIMSS site: <http://ustimss.msu.edu>; International TIMSS site: <http://timss.bc.edu/>). After years of ongoing science education reform, U.S. schools are now beginning to be held accountable for higher levels of performance among students.

The Move to High Stakes Testing

One prominent new strategy for ensuring accountability and higher performance among students has come to be known as *high-stakes testing*, the use of test scores to determine which students will graduate or which will be promoted from one grade to the next. In some cases the stakes may also include decisions about which teachers will get salary bonuses, or which schools will get extra funds to support academic improvements. This rapidly spreading practice was once described as "the latest silver bullet designed to cure all that ails public education" (Kunen, 1997). But is it a bullet that cures, or does it kill? Does high-stakes accountability testing support standards-based reform efforts, or hinder them?

While proponents see high-stakes testing as a means of holding schools, teachers, and students to high standards, some view testing as being inconsistent with the stated goals of the NSES (Huber & Moore, 2000). Indeed, the NSES (pp. 52, 72, 113, & 239) call for less emphasis on external assessments and standardized tests unrelated to *Standards-based* programs and practices.

Response to standardized tests by the general public seems mixed. According to the most recent Phi Delta Kappa/Gallup Poll. (Available online at: <http://www.pdkintl.org/kappan/k0109gal.htm>). Of those polled, 44% thought there was just the right amount of emphasis on standardized testing, but 51% of public school parents opposed "using a single standardized test ... to determine whether a student should be promoted from grade to grade." Interestingly, only 45% of public school parents opposed "using a single standardized test ... to determine

whether a student should receive a high school diploma."

Stronger support is provided by a survey sponsored by The Business Roundtable (Available online at: <http://www.brtable.org/press.cfm#53>). Indicating that 65% of parents and 70% of the general public support a policy of requiring students to "pass statewide tests before they can graduate from high school, even if they have passing grades in their classes." This is viewed as good news for the business community that has supported the push for rigorous education standards for some time.

Unintended Outcomes of High-Stakes Testing

Despite broad-based support for high-stakes testing, there is organized opposition (Schrag, 2000). Complaints range from concerns that the testing is "killing" innovative teaching and driving out good teachers to claims that tests overstress young students and are unfair to poor and minority students and others who lack test-taking skills. Others say that such tests limit the curriculum and "snuff out both creative teaching and the joy of learning" (Blair & Archer, 2001).

At a more fundamental level, questions about the validity of high-stakes tests and the ways they are being used and interpreted threaten to undermine the entire standards-based reform movement (Domenech, 2000). Objectivity and "teaching to the tests" are real concerns. In addition to narrowing the focus of instruction and assessment, there is an added risk of overburdening students and teachers through practices that may lead to inappropriate inferences about student performance (Ananda & Rabinowitz, 2000).

Finally, some claim that high-stakes testing creates a system that is unfair and destructive to learning, and that tougher standards and standardized testing are uniquely harmful to low-income and minority students (Kohn, 2000). While high-stakes testing may raise the level of education overall and raise the level of success by some students after graduation, the tests will exacerbate the problems of those already at risk or struggling to overcome disadvantaged backgrounds (Orfield & Kornhaber, 2001).

Status of Testing in Science

During Fall, 2001, the Council of Chief State School Officers (CCSSO) published the *1999-2000 Annual Survey of State Student Assessment Programs* (See http://publications.ccsso.org/ccsso/publication_detail.cfm?PID=350). Of states surveyed, 39 reported some form of proficiency testing in science being included in the state testing program. The results of state testing programs were used in making decisions about student promotion or retention in nine states, and passing scores were required for graduation in 17 states. Test results were included in reports of school performance in 37 states, and test results were used in making school improvement plans in 30 states. In only six states were test results used for staff accountability purposes, with four states using results as a basis for monetary rewards, such as bonuses.

The impact of one state testing program has been closely examined (Huber & Moore, 2000), and evidence indicates that

5866199

the highly publicized, model program has "derailed efforts to implement standards-based reforms" in science. Though high-stakes testing programs and the NSES appear to be at cross-purposes in several regards, two areas are of particular concern: equity and excellence.

With regard to equity issues, the testing program accentuates well-documented barriers to learning science among selected groups of students. In addition to evidence that the tests are biased (see Huber & Moore, 2000), they provide the basis for sanctions against the low-performing schools that are in need of most help in develop locally relevant programs.

Even if equity issues were adequately resolved, there remains a fundamental clash between high-stakes testing and the central features of the NSES. The NSES place great importance on learning through inquiry, de-emphasizing science as a body of factual knowledge to focus on science as a way of knowing. It is hoped that students will learn how to frame questions and use inquiry to find answers, investigating real problems. High-stakes standardized testing has the opposite thrust, focusing on a broad body of factual knowledge. May have claimed that this emphasis will pressure teachers to "teach to the test" and focus on particular subjects, and that appears to be happening. In a survey of teachers (Jones, Jones, Hardin, Chapman, Yarbrough, & Davis, 1999), 80% of participating teachers reported spending over 21% of their instructional time practicing for End-of-Grade tests, with over 28% of the teachers spending from 61% to 100% of their instructional time practicing for the tests.

Next Moves

It has been pointed out that assessment must be aligned with curriculum and instruction to support learning (Pellegrino, Chudowsky, and Glaser, 2001), so this is an issue that needs much attention as the practice of high-stakes testing spreads. Webb (1999) has described the development of new procedures for determining the degree of alignment of science and mathematics standards with assessment. Three states volunteered to have their science standards and assessments analyzed for two or three grade levels, and the results of analysis are highly variable. Four criteria were used in measuring the degree of alignment:

- *Categorical Coherence*—the extent to which the categories of content appear in both standards and assessment documents.
- *Depth-of Knowledge Consistency*—the extent to which the cognitive demand of tests reflects what students are expected to know.
- *Range-of Knowledge Correspondence*—the extent to which the span of knowledge required on the assessment matches the span of knowledge expected of students.
- *Balance of representation*—the extent to which test items are evenly distributed across objectives.

Though the results of this case study are not generalizable beyond the participating states, it is interesting to note the pattern of correspondence between science standards and assessments across the criteria. Though there was judged to be 100% alignment in terms of *Balance of representation*, there was little

Range-of Knowledge Correspondence (0% to 33%). Though somewhat better, the *Categorical Coherence* (38% to 67%) and *Depth-of Knowledge Consistency* (25% to 83%), ranged from poorly to highly aligned among individual states.

The most important outcome of the study is the emergence of a process to judge the alignment between science standards and assessments, and more states much carefully consider this issue. The CCSO has developed a research tool base on these results, the Surveys of Enacted Curriculum (SEC), that provides a practical, efficient means of obtaining consistent data on mathematics and science education practices through teacher reports. This approach enables schools, districts, or states to analyze current classroom practices in relation to content standards and facilitate program evaluations, curriculum improvements, interpretation of student assessment results, and alignment of curricula with standards (See <http://www.ccsso.org/sec.html>). It is imperative that states basing important decisions about students, teachers, and schools on high-stakes tests begin using or developing tools like this. States must quickly begin a process of alignment between standards and assessment so that "teaching to the test" becomes "teaching to the standards" in science.

References

- Ananda, S. & Rabinowitz, S. (2000). *The High Stakes of HIGH-STAKES Testing* (Policy Brief). San Francisco, CA: WestEd. [ED455254]
- Blair, J., & Archer, J. (2001, July 11). NEA members denounce high-stakes testing. *Education Week*, 20 (42). Web-only at http://www.edweek.org/ew/story.cfm?slug=42neatest_web_h20.
- Domenech, D. A. (2000, December). My Stakes Well Done. *School Administrator*, 57 (11), 16-19.
- Huber, R. A. & Moore, C. J. (2000). Educational reform through high stakes testing—Don't go there. *Science Educator*, 9 (1), 7-13.
- Jones, B. D., Jones, G. M., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (1999, November). The impact of high-stakes testing on teachers and students. *Phi Delta Kappan*, 199-203.
- Kohn, A. (2000, September-October). Burnt at the High Stakes. *Journal of Teacher Education*, 51 (4), 315-27.
- Kunen, J. S. (1997, June 16). The test of their lives. *Time*, 149 (24), 62-63.
- Miller, D. W. (2001, March). Scholars Say High-Stakes Tests Deserve a Failing Grade. *Chronicle of Higher Education*, 47 (25), A14-A16.
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press. (Available online at: <http://stills.nap.edu/html/nse/>)
- Orfield, G. & Kornhaber, M. L. (Eds.). (2001). *Raising standards or raising barriers?* Washington, DC: The Century Foundation Press.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press. (Available online at: <http://www.nap.edu/catalog/10019.html>).
- Schrag, P. (2000, August). "High stakes are for tomatoes." *The Atlantic Monthly*, 286 (2), 19-21. (Available online at: <http://www.theatlantic.com/issues/2000/08/schrag.htm>)
- Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states* (Research Monograph No. 18). Washington, DC: Council of Chief State School Officers).